Medical Science

Optimizing sample size estimation and statistical analysis in health sciences: A comprehensive guide using the S2SAC tool

Mohamad Arif Awang Nawi*

To Cite:
Nawi MAA. Optimizing sample size estimation and statistical analysis in health sciences: A comprehensive guide using the S2SAC tool. *Medical Science* 2024; 28: e75ms3380
doi: https://doi.org/10.54905/disssi.v28i149.e75ms3380

Authors' Affiliation:
Biostatistics Unit, School of Dental Sciences, Health Campus, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia

***Corresponding Author**
Biostatistics Unit, School of Dental Sciences, Health Campus, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan,
Malaysia
Email: mohamadarif@usm.my
ORCID: 0000-0002-4922-2881

Peer-Review History
Received: 30 April 2024
Reviewed & Revised: 04/May/2024 to 05/July/2024
Accepted: 09 July 2024
Published: 16 July 2024

Peer-review Method
External peer-review was done through double-blind method.

Medical Science
pISSN 2321–7359; eISSN 2321–7367

ABSTRACT

Statistical errors in health sciences research can greatly undermine the dependability of study results, negatively impacting clinical decision-making and policy development. This work focuses on the crucial task of minimizing potential statistical errors that arise from inadequate sample size estimation and improper analysis methodologies. The Sample Size & Statistical Analysis Calculator (S2SAC) is a tool specifically developed to improve the precision and dependability of statistical analyses in health research. S2SAC is a statistical software that can be easily used within the Microsoft Excel environment. It has a user-friendly interface and provides a wide range of statistical operations, such as outlier detection, normality testing, and both parametric and non-parametric tests. This application automates the intricate computations necessary for accurate sample size estimation and reliable statistical analysis, allowing researchers without substantial statistical experience to utilize advanced statistical approaches easily. This paper examines the potential of S2SAC through in-depth case studies, showcasing its efficacy in enhancing research design and safeguarding the credibility of research outcomes in the field of health sciences.

Keywords: Sample size estimation, Statistical analysis, Health sciences, S2SAC (Sample Size & Statistical Analysis Calculator), Statistical software

1. INTRODUCTION

Statistical analysis is crucial in the field of health sciences as it has a direct impact on decisions that influence patient care and public health policy. The accuracy and credibility of study conclusions depend on the utilization of precise and strong statistical methodologies. Improper use of these approaches can result in mistakes that could jeopardize patient outcomes and public health activities.

DISCOVERY SCIENTIFIC SOCIETY

Therefore, researchers must use sound statistical practices to ensure that conclusions drawn from data are robust and dependable (Kirkwood and Sterne, 2003; Altman and Bland, 2011). Choosing an inappropriate statistical test can lead to incorrect conclusions. For example, applying a statistical test like a t-test that assumes a normal distribution on data that do not meet this assumption can increase the likelihood of Type I errors, resulting in false positive findings. The error could result in the mistaken belief that a treatment is effective when it is not.

Improperly used tests undermine the credibility of scientific discoveries, potentially resulting in ineffective or hazardous medical procedures. Comprehending the attributes of the data and selecting the appropriate tests are essential for establishing the credibility of any research (Altman and Bland, 1991). Many health science researchers need access to readily comprehensible statistical methods specifically designed to address these issues. Researchers lacking extensive statistical expertise may encounter challenges while utilizing statistical software, as it typically requires a deep understanding of statistical theories and techniques (Field, 2013; Leek and Peng, 2015). Health science researchers encounter many statistical issues that can impact the integrity of their studies, including the determination of the optimal sample size. It is essential to achieve statistically meaningful results without expending money on large samples.

Failure to estimate the sample size accurately can result in studies that are unable to identify significant effects, while overestimating can needlessly increase the expenses and duration of research (Lenth, 2001). A study with an insufficient number of participants may need to be able to identify differences or effects that are truly present, resulting in Type II errors (false negatives). This error can lead to the early rejection of treatments that have the potential to be beneficial (Button et al., 2013). On the other hand, excessively large investigations not only squander resources but also unnecessarily expose more individuals to experimental treatments or procedures. According to Moore et al., (2011), this can be considered morally and financially indefensible. Estimating sample size accurately requires finding a balance between statistical power (the likelihood of detecting a true effect if it exists), the level of significance (the threshold for determining statistical significance), and the effect size (the expected magnitude of the effect) (Cohen, 1988).

The Sample Size & Statistical Analysis Calculator (S2SAC) is a versatile statistical tool developed specifically for use in the Microsoft Excel platform. The S2SAC platform offers a user-friendly interface that combines features for identifying outliers, testing for normality, and performing both parametric and non-parametric statistical tests. The purpose of this tool is to streamline the statistical analysis process by enabling researchers to do various analyses simultaneously, ranging from assessing data normality to executing the final statistical tests. This integration into Excel makes S2SAC accessible to a wide range of users, from students to experienced researchers, facilitating robust data analysis without the need for expensive or complex statistical software. The primary aim of this paper is to explore and elucidate the capabilities and benefits of the Sample Size & Statistical Analysis Calculator (S2SAC) as a comprehensive statistical tool integrated within the Microsoft Excel environment.

2. METHODOLOGY

The Sample Size & Statistical Analysis Calculator (S2SAC) is a statistical tool designed within the Microsoft Excel platform to provide a user-friendly interface for conducting comprehensive statistical analyses. S2SAC integrates various statistical functions such as outlier detection, normality testing, and both parametric and non-parametric statistical tests, aiming to simplify the statistical analysis process for researchers across different levels of expertise. In the data analysis process, outlier detection is a critical step that ensures the accuracy and reliability of statistical results. The Sample Size & Statistical Analysis Calculator (S2SAC) employs two robust methods for identifying outliers: Z scores and box plots. Specifically, Z scores are computed for each data point to quantify its deviation from the mean, represented in terms of standard deviations.

A common criterion used in statistical analyses is that any data point found to be more than three standard deviations from the mean is considered an outlier. This method helps pinpoint data points that are statistically significant anomalies. Alongside Z scores, box plots are also used within S2SAC to provide a visual representation of data distribution. Outliers are easily spotted in box plots as those points that lie beyond the plot's whiskers, which typically extend to 1.5 times the interquartile range from the quartiles (Tukey, 1977). This combination of Z scores and box plots allows researchers to accurately detect and assess outliers, thereby mitigating their potential to skew the results of the analysis. In the context of statistical analysis, especially when using parametric tests, the assumption of data normality is crucial.

The Sample Size & Statistical Analysis Calculator (S2SAC) integrates multiple approaches to evaluate this assumption, facilitating a thorough examination of data distributions. Among the methods it utilizes are the Shapiro-Wilk, Kolmogorov-Smirnov, and Anderson-Darling tests, each catering to different aspects of normality assessment. The Shapiro-Wilk test is often preferred for smaller sample sizes due to its sensitivity (Razali and Wah, 2011). The Kolmogorov-Smirnov test offers a non-parametric option that contrasts the empirical distribution of a sample with a reference distribution. However, the Anderson-Darling test is particularly noted for its effectiveness in detecting deviations from normality in the data tails, making it invaluable for comprehensive data analysis (Stephens, 1974).

This test modifies the Kolmogorov-Smirnov test by giving more weight to the tails where deviations are often critical yet hard to detect. The integration of these tests into S2SAC allows researchers to confidently assess normality, ensuring that the prerequisites for parametric testing are rigorously evaluated. In statistical analysis, choosing between parametric and non-parametric tests depends crucially on the data type and distribution. The Sample Size & Statistical Analysis Calculator (S2SAC) provides parametric tests such as the one-sample t-test, paired-sample t-test, independent-sample t-test, and one-way ANOVA, which are suitable for data with a normal distribution and are used to analyze means and variances. Additionally, it includes the Pearson correlation to evaluate linear relationships between variables (Rodgers and Nicewander, 1988).

For data that do not meet these normality assumptions, S2SAC offers non-parametric methods like the Wilcoxon tests for comparing medians, the Mann-Whitney U test for independent samples, the Kruskal-Wallis test for more than two groups, and the Spearman correlation for assessing non-linear relationships (Corder and Foreman, 2014). This dual approach allows for flexible and accurate handling of diverse data types. These comprehensive testing options within S2SAC ensure that researchers can select the most appropriate statistical tests based on the characteristics of their data, thereby enhancing the robustness and validity of their research findings.

Sample size estimation is a critical component of research design that S2SAC addresses comprehensively. This feature in S2SAC allows researchers to calculate the necessary sample sizes required to achieve adequate statistical power in various testing scenarios (Figure 1). These calculations are grounded in the principles of statistical hypothesis testing, where power—the probability of rejecting the null hypothesis when it is false—depends significantly on several factors: the expected effect sizes, standard deviations of the data, alpha levels, and the desired power of the test. Typically, alpha levels are set at 0.05 to maintain a 95% confidence level, reflecting a 5% chance of committing a Type I error. In contrast, the power is commonly set between 80% and 90%, depending on the rigour required by the research (Cohen, 1988).

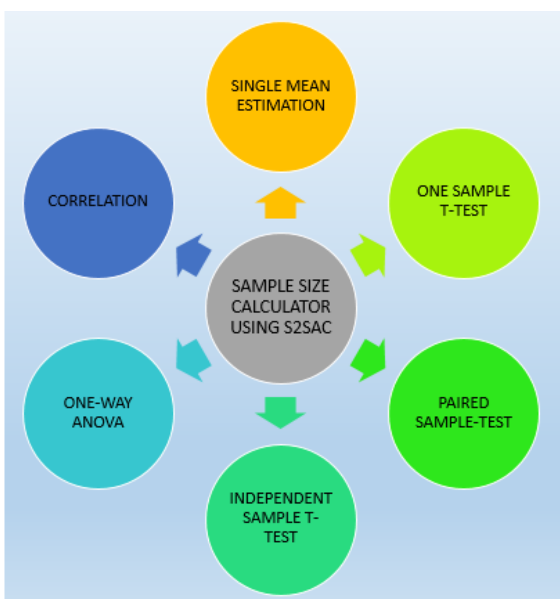

Figure 1 Sample Size Calculation based on Statistical Analysis using S2SAC

S2SAC's versatility is evident as it supports sample size estimations for a variety of analyses. This calculator includes comparisons of means, which might use t-tests or ANOVAs; correlation analyses, which assess the strength of relationships between variables using Pearson or Spearman correlation coefficients; and analyses of rates and proportions, which often involve binomial data. Each type of analysis may require different considerations for estimating sample size, such as the expected differences in means or proportions, the variability of the data, or the strength of a correlation. By integrating these functionalities into a user-friendly tool, S2SAC significantly simplifies the complex process of planning for adequate statistical power, making robust and scientifically sound research more accessible to a wide range of investigators.

The methodology section of the research on the Sample Size & Statistical Analysis Calculator (S2SAC) elaborates on the design considerations pertinent to various types of studies that can leverage this tool effectively. S2SAC is adept at accommodating the unique demands of diverse research formats such as clinical trials, observational studies, and laboratory experiments. Its flexibility and adaptability make it an invaluable asset across different research domains. For clinical trials, S2SAC aids in the rigorous statistical planning needed to ensure reliable and valid results, which are crucial for regulatory approval and clinical applicability.

In observational studies, where the control over variables is less stringent, S2SAC's robust data handling and analysis capabilities help in drawing significant inferences from naturalistic data sets. Similarly, for laboratory experiments that often require precise measurements and quick data analyses, S2SAC streamlines the process, allowing researchers to focus more on experimental design and less on the complexities of data analysis. This adaptability underscores the tool's capacity to enhance research efficiency and effectiveness, thereby contributing to more scientifically sound conclusions.

3. RESULTS

S2SAC data can be evaluated statistically very easily, and running the data simultaneously (3 in 1) involves checking outliers, checking normality, and then conducting statistical analysis. This paper presents several case studies showing how S2SAC was used in health sciences research. S2SAC outcomes were examined in detail and compared with conventional techniques (Figure 2).

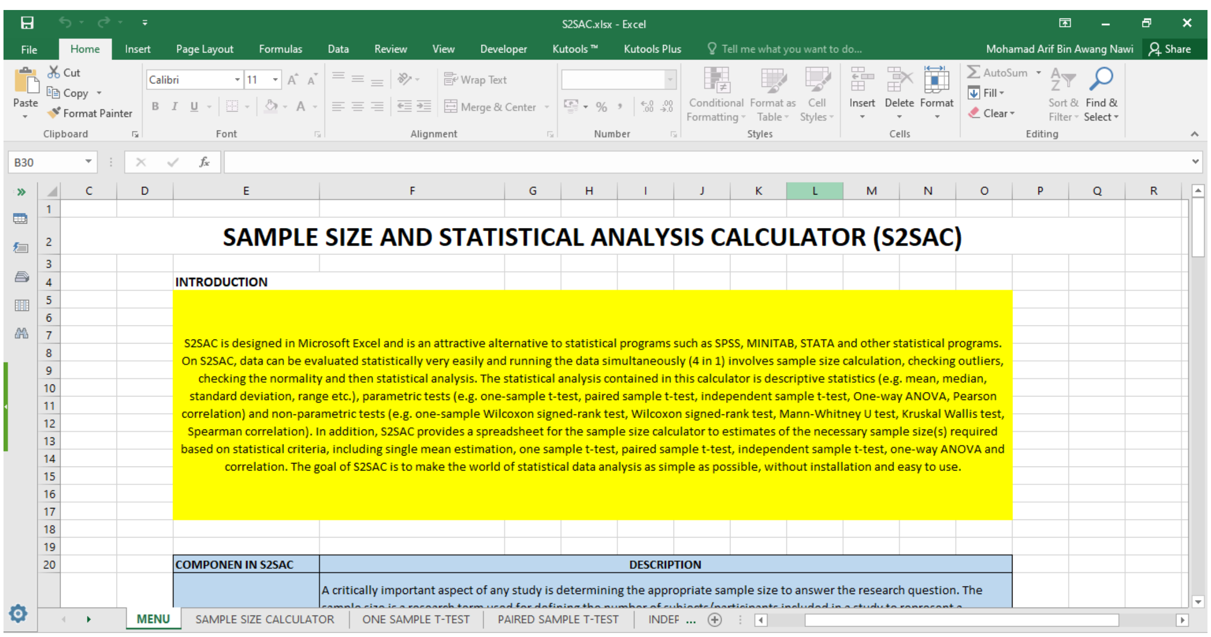

Figure 2 Interface of S2SAC Program Using Excel

Basic Steps Toward Sample Size Calculation

Step One: Objective

A clear and quantitatively achievable research objective is paramount in sample size estimation. The objective should define measurable or countable outcomes and predictors relevant to the research question. For instance, objectives such as determining the mean systolic blood pressure among specific cohorts or assessing the prevalence of a particular condition among defined populations

provide clear targets for sample size calculation. General objectives should be refined into specific, quantifiable components to facilitate accurate sample size determination.

Step Two: Hypothesis Testing or Estimation

Once the research objective is clarified, researchers must ascertain whether the objective necessitates hypothesis testing or estimation. For hypothesis testing, objectives typically involve comparing outcomes between groups or assessing relationships between variables, requiring the use of appropriate statistical tests such as t-tests or ANOVA. Estimation objectives, on the other hand, focus on determining population parameters with specified confidence intervals devoid of predictive variables. Clarifying the nature of the research objective is essential for selecting the appropriate statistical approach and sample size calculator.

Step Three: Sample Size Calculator

Following the delineation of the research objective and statistical approach, researchers can identify the suitable sample size calculator within S2SAC. By matching the research objective to the corresponding calculator, researchers can accurately estimate the sample size required for their study. For instance, objectives involving comparisons between group means may necessitate the use of calculators tailored to independent sample t-tests or ANOVA. In contrast, objectives centred on estimating population parameters may utilize calculators for single mean or proportion estimations.

Case Study I for Single Mean Estimation

For a single mean, the objective of a study is to estimate the mean of an outcome of interest in a population from data obtained from the sample, of which the outcome is measured on a numerical continuous scale. The calculator for single mean estimation is as follows (Figure 3):

SINGLE MEAN ESTIMATION	
Significance level (α)	
Standard deviation (σ)	
Precision (d)	
Drop-out	
Sample size	**#NUM!**
Sample size (with drop-out)	**#NUM!**

Indicator	**Explanation**
Significance level (α)	Researcher decide significant level used in the study. - If 95% confidence level, significance level (α) is 0.05 (Recommended value). - If 99% confidence level, significance level (α) is 0.01.
Standard deviation	Take from - Published papers in a similar study. If standard deviation (SD) value comes from the published paper, state the references in details. - Pilot study.
Precision/Margin of Error	Decide the precision or margin of error (Δ) based on clinical or biological reason.
Drop-out	Decide drop-out rate between 10%-30%.

Figure 3 Calculator, Indicator and Explanation for Single Mean Estimation

Case Study I

One researcher wants to study the level of fasting blood sugar in a group of elderly people in Malaysia. From previous research, the mean fasting blood sugar among 400 elderly people in Malaysia is 5.8 mmol/dl, and the standard deviation is 0.2 mmol/dl. He wishes to estimate the margin of error within 0.02 mmol/dl with 95% confidence level (so α = 0.05). How many sample sizes of the elderly in Malaysia should be included in this study? The minimum sample size required is 385 elderly people in Malaysia (Figure 4). If a 90% response rate is expected, a 428-person sample of elderly people in Malaysia is needed for this study.

SINGLE MEAN ESTIMATION	
Significance level (α)	0.05
Standard deviation (σ)	0.2
Precision (d)	0.02
Drop-out	10%
Sample size	**385**
Sample size (with drop-out)	**428**

Figure 4 Example of Single Mean Calculation for Case Study I

Case Study II for One-Sample T-Test

Sample size estimation for a one-sample t-test calculator to determine the number of samples where the objective of a study is to compare the mean of a population with a "test value" (hypothesized value of the mean in the population). The outcome is measured on a numerical continuous scale. The sample size estimation for the one-sample t-test calculator is as follows (Figure 5):

Case Study II

A researcher wants to compare the systolic blood pressure between patients at HUSM and the "test value" (hypothesized value of population mean). He tested each patient on a specific hypertension index, where anyone reaching a score of 120 mmHg (test value based on a trusted research organization) was considered to have a 'normal' systolic blood pressure level. How many patients at HUSM should be enrolled in the study? The investigator plans to use a 95% confidence level (so α = 0.05) and 80% power. The mean and standard deviation of systolic blood pressure were unknown. However, the investigators conducted a literature search and found the mean and standard deviation of systolic blood pressure to be 130mmHg (20mmHg).

Significance level (α) = 0.05

Power of Study = 80% = 0.8

Expected difference (Δ) = Sample mean - test value = 128mmHG – 120mmHg = 8 mmHg

Standard deviation (σ) of SBP from literature search = 20 mmHg

Drop-out between 10%-30%. The researcher decides 20%.

Based on Figure 6, the minimum sample size required is up to 50 patients. If the response rate is estimated at only 80%, then add another 20% for the data collection stage. Therefore, the sample needed in the study is 63 patients at HUSM.

ONE SAMPLE T-TEST	
Significance level (α)	
Power (1-β)	
Expected difference (Δ)	
Standard deviation (σ)	
Drop-out	
Sample size	**#DIV/0!**
Sample size (with drop-out)	**#DIV/0!**

Indicator	**Explanation**
Significance level (α)	Researcher decide significant level used in the study. - If 95% confidence level, significance level (α) is 0.05 (*Recommended value*). - If 99% confidence level, significance level (α) is 0.01.
Power (1-β)	Power, probability of making correct decision based on true real situation. Usually set at 80% (or 0.8) or 90% (or 0.9)
Expected difference (Δ)	Difference between the sample mean (based on expert opinion/previous study) and the "test value" (based on literature review/a trusted research organization/legal requirements/industry standards)
Standard deviation (σ)	The standard deviation of sample (based on previous study / pilot study).
Drop-out	Decide drop-out rate between 10%-30%.

Figure 5 Calculator, Indicator and Explanation for One Sample T-Test

ONE SAMPLE T-TEST	
Significance level (α)	0.05
Power (1-β)	0.8
Expected difference (Δ)	8
Standard deviation (σ)	20
Drop-out	20%
Sample size	**50**
Sample size (with drop-out)	**63**

Figure 6 Example of One Sample T-Test Calculation for Case Study II

Case Study III for Paired Sample T-Test

Sample Size estimation for Paired Samples t-tests is more complex than sample size calculation for the independent samples t-test. While the sample size requirement is smaller because the two samples are related or correlated, the calculation is somewhat complicated. The sample size is used to compare the means between two groups of populations, which are dependent, paired, or

matched. Sometimes, two groups are matched or paired. They are no longer independent. The usual statistical test is a paired sample t-test. The sample size calculator for the Paired sample t-test is as below (Figure 7):

PAIRED SAMPLE-TEST	
Significance level (α)	
Power (1-β)	
Expected difference (Δ)	
*Standard deviation of difference(σ)	
Drop-out	
Sample size	**#DIV/0!**
Sample size (with drop-out)	**#DIV/0!**

Indicator	**Explanation**
Significance level (α)	Researcher decide significant level used in the study. - If 95% confidence level, significance level (α) is 0.05 (*Recommended value*). - If 99% confidence level, significance level (α) is 0.01.
Power (1-β)	Power, probability of making correct decision based on true real situation. Usually set at 80% (or 0.8) or 90% (or 0.9)
Expected difference (Δ)	The absolute value of the hypothesized mean difference between two samples are related or correlated (based on expert opinion/literature research).
Standard deviation of difference (σ)	hypothesized standard deviation of differences (known for example from a Paired samples t-test from previous studies, or from the literature).
Drop-out	Decide drop-out rate between 10%-30%.

Figure 7 Calculator, Indicator and Explanation for Paired Sample-Test

Case Study III

A group of medical researchers wishes to determine the sample size needed if they would like to prove that a group of children in a community changes their body weight by at least 10 kg after a special 3-year diet intervention. The pre-intervention group and the post-intervention group are the same children. The literature research found that the standard deviation of the mean change in body weight was 20 kg. They choose a 95% confidence level (α = 0.05) and power 90%. How many children in a community should be enrolled in the study?

Significance level (α) = 0.05

Power of Study = 90% = 0.9

Expected difference (Δ) = The true difference in the mean effect of pre-intervention and post-intervention = at least 10 kg

The standard deviation of difference (σ) = Previously the standard deviation of the mean change in the body weight was 20 kg

Drop-out between 10%-30%. The researcher decides 20%.

The minimum sample size required is up to 43 children (Figure 8). If the response rate is estimated at only 80%, then add another 20% for the data collection stage. Therefore, the sample needed in the study is 54 children in a community.

PAIRED SAMPLE-TEST	
Significance level (α)	0.05
Power (1-β)	0.9
Expected difference (Δ)	10
*Standard deviation of difference(σ)	20
Drop-out	20%
Sample size	**43**
Sample size (with drop-out)	**54**

Figure 8 Example of Paired Sample-Test Calculation for Case Study III

Case Study IV for Independent Sample T-Test

Sample Size estimation for independent sample t-tests calculator to determine the number of samples where the objective of a study is to compare the means of a continuous outcome variable in two independent groups. The sample size calculator for independent sample t-tests is as below (Figure 9):

INDEPENDENT SAMPLE T-TEST	
Significance level (α)	
Power (1-β)	
Expected difference (Δ)	
Standard deviation (σ)	
Group Size Ratio (m)	
Drop-out	
Sample size per group	**#DIV/0!**
Sample size per group (with drop-out)	**#DIV/0!**

Indicator	**Explanation**
Significance level (α)	Researcher decide significant level used in the study. - If 95% confidence level, significance level (α) is 0.05 (*Recommended value*). - If 99% confidence level, significance level (α) is 0.01.
Power (1-β)	Power, probability of making correct decision based on true real situation. Usually set at 80% (or 0.8) or 90% (or 0.9)
Expected difference (Δ)	The absolute value of the hypothesized mean difference between two groups are independent (based on expert opinion/literature research).
Standard deviation (σ)	The within group standard deviation (based on previous study / pilot study).
Group Size Ratio (m)	The ratio between the two groups of subjects/patients/respondents.
Drop-out	Decide drop-out rate between 10%-30%.

Figure 9 Calculator, Indicator and Explanation for Independent Sample T-Test

Case Study IV

A medical lecturer from PPSG wants to calculate the sample size of a study that compares the mean diastolic blood pressure of treated and untreated hypertensive patients in a staff clinic, HUSM. He knows from literature elsewhere the biggest standard deviation of mean diastolic blood pressure in this type of patient was 10 mmHg. The difference in the mean diastolic blood pressure between treated and untreated hypertensive patients in the staff clinic, HUSM, is 5 mmHg. He would like to include the same number of patients in both groups and choose a 95% confidence level (α = 0.05) with a power of 90%.

Significance level (α) = 0.05

Power of Study = 90% = 0.9

Expected difference (Δ) = The difference mean of diastolic blood pressure between treated and untreated group = 5 mmHg

Standard deviation (σ) = From literature research, the biggest standard deviation of mean diastolic blood pressure = 10 mmHg

Group Size Ratio (m) = A medical lecturer would like to include the same number of patients in both groups = 1

Drop-out between 10%-30%. The researcher decides 10%.

The total number of sample size needed is 85 for each group of patients (Figure 10). If the response rate is estimated at only 90%, then add another 10% for the data collection stage. Therefore, the sample needed in the study is 95 patients for each group. Therefore, the total sample size needed in this study is 190 patients (treated and untreated hypertensive patients).

INDEPENDENT SAMPLE T-TEST	
Significance level (α)	0.05
Power (1-β)	0.9
Expected difference (Δ)	5
Standard deviation (σ)	10
Group Size Ratio (m)	1
Drop-out	10%
Sample size per group	**85**
Sample size per group (with drop-out)	**95**

Figure 10 Example of Independent Sample T-Test Calculation for Case Study IV

Case Study V for One-Way ANOVA

Sample size estimation for a one-way ANOVA calculator is used to determine the number of samples where the objective of a study is to compare the means of a continuous outcome variable in three or more independent groups. The sample size calculator for One-Way ANOVA is as below (Figure 11):

Case Study V

A dental researcher wants to compare the effect of different modes of tooth brushing education (lecture, video and pamphlet) on the dental plaque index (PI) of adolescents. The investigator measures the effectiveness of the different modes of tooth brushing education using a dental plaque (PI) score. The mean PI scores will classify as less than 0.1 (absence of plaque or perfect plaque control), between 0.1 to 1.0 (small amount of plaque or good plaque control), between 1.0 to 2.0 (average value or fair plaque control), and between 2.1 to 3 (bad plaque control). By considering at least 80% power and α = 0.05, how many adolescents are required in the education via lecture group, video group, and pamphlet groups? He knows from literature research elsewhere that the mean (SD) PI scores for the education via lecture group is 0.7 (0.4), education via video group is 0.95 (0.38) and education via pamphlet is 1.02 (0.39). He would like to include an equal number of patients in each group.

ONE-WAY ANOVA	
Significance level (α)	
Power (1-β)	
Expected difference (Δ)	
Standard deviation (σ)	
Number of Groups (must be in range 2 to 10)	
Group Size Ratio (m)	
Drop-out	
Sample size per group	**#DIV/0!**
Sample size per group (with drop-out)	**#DIV/0!**

Indicator	**Explanation**
Significance level (α)	Researcher decide significant level used in the study. - If 95% confidence level, significance level (α) is 0.05 (*Recommended value*). - If 99% confidence level, significance level (α) is 0.01.
Power (1-β)	Power, probability of making correct decision based on true real situation. Usually set at 80% (or 0.8) or 90% (or 0.9)
Expected difference (Δ)	The absolute value of the hypothesized mean difference between highest and smallest means (based on expert opinion/literature research).
Standard deviation (σ)	The within group standard deviation (SD) (based on previous study / pilot study).
Number of Groups	Total number of group in study (must be in range 2 to 10).
Group Size Ratio (m)	The ratio between the groups of subjects/patients/respondents.

Figure 11 Calculator, Indicator and Explanation for One-Way ANOVA

Significance level (α) = 0.05
Power of Study = 80% = 0.8
Expected difference (Δ) = From literature research, the mean difference between the highest and smallest means of PI score = 1.02 – 0.7 = 0.32
Standard deviation (σ) = From literature research, the biggest standard deviation of mean PI score = 0.4
Number of groups = Total number of groups in the study = 3
Group Size Ratio (m) = A dental researcher would like to include an equal number of adolescents in the groups = 1
Drop-out between 10%-30%. The researcher decides 10%.

The sample size needed is 31 for each group of adolescents (Figure 12). If the response rate is estimated at only 90%, then add another 10% for the data collection stage. So, the sample needed in the study is 35 adolescents for each group. Therefore, the total sample size needed in this study is 35 adolescents' x 3 (number of groups) = 105 adolescents.

ONE-WAY ANOVA	
Significance level (α)	0.05
Power (1-β)	0.8
Expected difference (Δ)	0.32
Standard deviation (σ)	0.4
Number of Groups (must be in range 2 to 10)	3
Group Size Ratio (m)	1
Drop-out	10%
Sample size per group	**31**
Sample size per group (with drop-out)	**35**

Figure 12 Example of One-Way ANOVA Calculation for Case Study V

Case Study VI for Pearson/Spearman Correlation

Sample size estimation for correlation analysis should be in line with the study objective. Sample Size estimation for correlation calculator to determine the number of samples where the objective of a study is to measure the strength of a linear association between two variables. The sample size calculator for the Pearson/Spearman correlation is as below (Figure 13):

Case Study VI

A researcher at Hospital Universiti Sains Malaysia aims to examine the relationship between systolic blood pressure levels and body mass index (BMI) readings among patients. Previous research indicated a correlation coefficient of 0.53 between these two variables. A two-sided hypothesis test will be conducted at α = 0.05 to determine if there is a significant relationship between systolic blood pressure and BMI. How many patients should be included in the study to ensure that the test has 90% power to detect a correlation coefficient of at least 0.53?

Significance level (α) = 0.05

Power of Study = 90% = 0.9

Expected correlation coefficient, r = The correlation coefficient value between systolic blood pressure level and body mass index (BMI) = 0.53

Drop-out between 10%-30%. The researcher decides 10%.

The minimum number of samples needed to detect at least a correlation coefficient of 0.53 is 34 patients (Figure 14). If the response rate is estimated at only 90%, then add another 10% for the data collection stage. Therefore, the total sample needed in the study is 38 patients.

PEARSON/SPEARMAN CORRELATION	
Significance level (α)	
Power (1-β)	
Expected correlation coefficient, *r*	
Drop-out	
Sample size	**#NUM!**
Sample size (with drop-out)	**#NUM!**

Indicator	**Explanation**
Significance level (α)	Researcher decide significant level used in the study. - If 95% confidence level, significance level (α) is 0.05 (*Recommended value*). - If 99% confidence level, significance level (α) is 0.01.
Power (1-β)	Power, probability of making correct decision based on true real situation. Usually set at 80% (or 0.8) or 90% (or 0.9)
Expected correlation coefficient, r	The correlation coefficient value between two numerical variables (based on previous/ literature research).
Drop-out	Decide drop-out rate between 10%-30%.

Figure 13 Calculator, Indicator and Explanation for Pearson/Spearman Correlation

PEARSON/SPEARMAN CORRELATION	
Significance level (α)	0.05
Power (1-β)	0.9
Expected correlation coefficient, *r*	0.53
Drop-out	10%
Sample size	**34**
Sample size (with drop-out)	**38**

Figure 14 Example of Pearson/Spearman Correlation Calculation for Case Study VI

4. DISCUSSION

The Sample Size & Statistical Analysis Calculator (S2SAC) demonstrated significant utility in simplifying and enhancing the statistical analysis processes for health science research. By automating complex computations required for accurate sample size estimation and statistical analysis, S2SAC addresses a crucial need in the field where statistical errors can compromise the reliability of research findings (Kirkwood and Sterne, 2003; Altman and Bland, 2011). The user-friendly interface and integration within the Microsoft Excel environment make advanced statistical methods accessible to researchers with varying levels of statistical expertise. Health science research often encounters statistical challenges such as improper sample size estimation, which can lead to Type I and Type II errors, thereby affecting the validity of the study outcomes (Button et al., 2013).

The S2SAC tool's ability to precisely estimate sample sizes based on desired power and significance levels ensures that studies are adequately powered to detect meaningful effects without unnecessary resource expenditure (Lenth, 2001). This balance between statistical power, significance level, and effect size is crucial for robust research design (Cohen, 1988). The comprehensive range of statistical tests available in S2SAC, including parametric and non-parametric options, allows researchers to choose appropriate tests based on data characteristics. This flexibility is critical for maintaining the credibility of scientific discoveries, as improper test selection can lead to incorrect conclusions and potentially hazardous recommendations (Altman and Bland, 1991).

By incorporating robust methods for outlier detection and normality testing, S2SAC ensures that the underlying assumptions of statistical tests are rigorously evaluated, further enhancing the reliability of research findings (Razali and Wah, 2011; Stephens, 1974). The practical applications of S2SAC were illustrated through various case studies, showcasing its efficacy in different research scenarios. For instance, the tool's ability to accurately calculate sample sizes for studies involving single mean estimation, one-sample t-tests, paired sample t-tests, independent sample t-tests, one-way ANOVA, and correlation analyses demonstrates its versatility (Moore et al., 2011). These case studies highlight the tool's potential to streamline the research process, from planning and design to data analysis and interpretation.

5. CONCLUSION

Conclusion

The Sample Size and Statistical Analysis Calculator (S2SAC) proves to be an invaluable resource for educational researchers, providing a comprehensive array of tools for sample size estimation and statistical analysis. Through the development and application of S2SAC, this research endeavours to streamline and enhance the methodological rigour of educational research endeavours. S2SAC's user-friendly interface and diverse array of sample size calculators empower researchers to make informed decisions regarding sample size requirements, ensuring studies are adequately powered to detect meaningful effects. By automating the calculation process and adhering to established guidelines and best practices, S2SAC promotes transparency and reproducibility in educational research practices. While S2SAC presents numerous benefits, including efficiency, accuracy, and standardization, it is not without limitations. Researchers must remain vigilant in considering the assumptions underlying sample size calculations and interpreting statistical outputs accurately.

Moreover, the generalizability of findings obtained using S2SAC may be contingent on sample characteristics and contextual factors unique to educational settings. Looking ahead, continued refinement and validation of S2SAC algorithms, coupled with efforts to enhance user accessibility and training resources, are essential for promoting its widespread adoption and usability among educational researchers. By leveraging S2SAC's capabilities, researchers can advance methodological rigour, transparency, and reproducibility in educational research, ultimately contributing to evidence-based decision-making and policy development in the field. In conclusion, S2SAC stands as a testament to the ongoing evolution of statistical tools in educational research, empowering researchers to navigate the complexities of sample size determination and statistical analysis with confidence and precision.

Benefits of S2SAC in Educational Research

The integration of S2SAC into educational research practices presents numerous benefits. Firstly, S2SAC enhances the efficiency and accuracy of sample size estimation, ensuring studies are adequately powered to detect meaningful effects. By automating the calculation process and providing user-friendly interfaces, S2SAC empowers researchers to make informed decisions regarding sample size requirements based on their research objectives and statistical criteria. Moreover, the availability of diverse sample size calculators within S2SAC enables researchers to address a wide range of research questions and experimental designs, from simple comparisons of means to complex multivariate analyses. Furthermore, S2SAC promotes transparency and reproducibility in educational research by providing a standardized framework for sample size determination and statistical analysis. By adhering to established guidelines and best practices embedded within S2SAC, researchers can enhance the credibility and reliability of their findings, thereby contributing to the advancement of knowledge in the field.

Limitations and Considerations

Despite its utility, S2SAC has certain limitations that warrant consideration. Firstly, the accuracy of sample size estimates generated by S2SAC relies on several assumptions, encompassing the normality of data distributions and the homogeneity of variances. Deviations from these assumptions may impact the validity of sample size calculations and subsequent statistical analyses. Additionally, while S2SAC provides a comprehensive suite of statistical tools, its effectiveness is contingent on researchers' proficiency in understanding and interpreting statistical outputs. Misinterpretation of results or inappropriate application of statistical tests may compromise the validity of research findings. Moreover, the generalizability of findings obtained using S2SAC may be influenced by sample characteristics and contextual factors unique to educational settings. Researchers should exercise caution when extrapolating results from S2SAC analyses to broader populations or educational contexts, considering the potential influence of confounding variables and sampling biases.

Future Directions

There are several avenues for enhancing the utility and impact of S2SAC in educational research. Firstly, continued refinement and validation of sample size estimation algorithms within S2SAC are necessary to accommodate diverse research designs and analytical techniques. Collaboration between statisticians, software developers, and educational researchers is essential for incorporating advanced statistical methodologies and addressing emerging research challenges. Furthermore, efforts to enhance user accessibility and training resources for S2SAC are paramount in promoting its widespread adoption and usability among educational researchers.

Training workshops, online tutorials, and user manuals tailored to researchers' skill levels can empower users to leverage the full potential of S2SAC in their research endeavours. In conclusion, the integration of S2SAC into educational research practices holds immense promise for advancing methodological rigour, transparency, and reproducibility in the field. By providing researchers with a powerful toolkit for sample size estimation and statistical analysis, S2SAC contributes to the generation of high-quality evidence that informs evidence-based decision-making and policy development in education.

Acknowledgments

The authors would like to express their gratitude to the Ministry of Higher Education (MOHE) Malaysia for providing the research funding Fundamental Research Grant Scheme (FRGS) (FRGS/1/2023/STG06/USM/03/3).

Author Contributions

Mohamad Arif Awang Nawi: Contributed in the drafting of the manuscript, conception and design of the study, analysis, interpretation of data and final approval.

Informed Consent

Not applicable.

Funding

This study was funded by Ministry of Higher Education (MOHE) Malaysia for providing the research funding Fundamental Research Grant Scheme (FRGS) (FRGS/1/2023/STG06/USM/03/3).

Conflict of interest

The authors declare that there is no conflict of interests.

Data and materials availability

All data sets collected during this study are available upon reasonable request from the corresponding author.

REFERENCES

1. Altman DG, Bland JM. How to obtain the P value from a confidence interval. BMJ 2011; 343:d2304. doi: 10.1136/bmj.d2304
2. Altman DG, Bland JM. Improving doctors' understanding of statistics. J R Stat Soc Series A Stat Soc 1991; 154(2):223-267. doi: 10.2307/2983040
3. Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci. 2013;14(5):365-76. doi: 10.1038/nrn3475. Erratum in: Nat Rev Neurosci 2013; 14(6):451
4. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Hillsdale NJ: Lawrence Erlbaum Associates, Publishers, 1988.
5. Corder GW, Foreman DI. Nonparametric Statistics: A Step-by-Step Approach. John Wiley & Sons, 2014.
6. Field A. Discovering Statistics Using IBM SPSS Statistics: And Sex and Drugs and Rock "N" Roll. 4th Edition, Sage, Los Angeles, London, New Delhi, 2013.
7. Kirkwood BR, Sterne JAC. Essential Medical Statistics. 2nd Edition, Blackwell Science, Oxford, 2003; 420-421.
8. Leek J, Peng RD. Statistics: P values are just the tip of the iceberg. Nature 2015; 520(7549):612. doi: 10.1038/520612a
9. Lenth RV. Some practical guidelines for effective sample size determination. Am Stat 2001; 55(3):187-93.
10. Moore DS, McCabe GP, Craig BA. Introduction to the Practice of Statistics. New York: WH Freeman and Company; 2011.
11. Razali NM, Wah YB. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. J Stat Model Anal 2011; 22:21-33.
12. Rodgers JL, Nicewander WA. Thirteen ways to look at the correlation coefficient. Am Stat 1988; 42(1):59-66.
13. Stephens MA. EDF Statistics for Goodness of Fit and Some Comparisons. J Am Stat Assoc 1974; 69(347):730-737.